Three.V Change of Basis

Linear Algebra

Jim Hefferon

http://joshua.smcvt.edu/linearalgebra

Changing representations of vectors

Coordinates vary with the bases

Example Consider this vector $\vec{v} \in \mathbb{R}^3$ and bases for the space.

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \mathcal{E}_3 = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle \quad B = \langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle$$

With respect to the different bases, the coordinates of $\vec{v}$ are different.

$$\mathrm{Rep}_{\mathcal{E}_3}(\vec{v}) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \mathrm{Rep}_{B}(\vec{v}) = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

Here we will see how to convert between the two representations: given two bases for a space we want a formula that converts the representation of a vector with respect to the first basis to the representation with respect to the second.

Change of basis matrix

If we think of translating from $\mathrm{Rep}_{\mathrm{B}}(\vec{v})$ to $\mathrm{Rep}_{\mathrm{D}}(\vec{v})$ as holding the vector constant then this is the appropriate arrow diagram.

$$\begin{array}{c} V_{\textit{wrt}\ \mathrm{B}} \\ \mathrm{id}\big\downarrow \\ V_{\textit{wrt}\ \mathrm{D}} \end{array}$$

(This diagram is vertical to fit with the ones in the next subsection.)

Change of basis matrix

If we think of translating from $\operatorname{Rep}_{B}(\vec{v})$ to $\operatorname{Rep}_{D}(\vec{v})$ as holding the vector constant then this is the appropriate arrow diagram.

$$\begin{tikzcd} V_{wrt\ B} \arrow[d, "\mathrm{id}"'] \\ V_{wrt\ D} \end{tikzcd}$$

(This diagram is vertical to fit with the ones in the next subsection.)

1.1 *Definition* The *change of basis matrix* for bases $B, D \subset V$ is the representation of the identity map $\mathrm{id}\colon V \to V$ with respect to those bases.

$$\operatorname{Rep}_{B,D}(\mathrm{id}) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \operatorname{Rep}_{D}(\vec{\beta}_1) & \cdots & \operatorname{Rep}_{D}(\vec{\beta}_n) \\ \vdots & & \vdots \end{array}\right)$$

1.3 *Lemma* Left-multiplication by the change of basis matrix for B, D converts a representation with respect to B to one with respect to D. Conversely, if left-multiplication by a matrix changes bases $M \cdot \mathrm{Rep}_B(\vec{v}) = \mathrm{Rep}_D(\vec{v})$ then M is a change of basis matrix.

Proof The first sentence holds because matrix-vector multiplication represents a map application and so $\mathrm{Rep}_{B,D}(\mathrm{id}) \cdot \mathrm{Rep}_B(\vec{v}) = \mathrm{Rep}_D(\,\mathrm{id}(\vec{v})\,) = \mathrm{Rep}_D(\vec{v})$ for each $\vec{v}$. For the second sentence, with respect to B, D the matrix M represents a linear map whose action is to map each vector to itself, and is therefore the identity map. QED

Example Two bases for $\mathcal{P}_2$ are $B = \langle 1, 1+x, 1+x+x^2 \rangle$ and $D = \langle x^2-1, x, x^2+1 \rangle$. Compute $\operatorname{Rep}_{B,D}(\mathrm{id})$ in the same way that we compute the representation of any function: find $\operatorname{Rep}_D(\mathrm{id}(1))$, $\operatorname{Rep}_D(\mathrm{id}(1+x))$, and $\operatorname{Rep}_D(\mathrm{id}(1+x+x^2))$.

$$\operatorname{Rep}_D(1) = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \end{pmatrix} \quad \operatorname{Rep}_D(1+x) = \begin{pmatrix} -1/2 \\ 1 \\ 1/2 \end{pmatrix} \quad \operatorname{Rep}_D(1+x+x^2) = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

We put them together into the change of basis matrix.

$$\begin{pmatrix} -1/2 & -1/2 & 0 \\ 0 & 1 & 1 \\ 1/2 & 1/2 & 1 \end{pmatrix}$$

For instance, we have this.

$$\operatorname{Rep}_B(2-x+3x^2) = \begin{pmatrix} 3 \\ -4 \\ 3 \end{pmatrix} \qquad \operatorname{Rep}_D(2-x+3x^2) = \begin{pmatrix} 1/2 \\ -1 \\ 5/2 \end{pmatrix}$$

The change of basis matrix does indeed do the conversion.

$$\begin{pmatrix} -1/2 & -1/2 & 0 \\ 0 & 1 & 1 \\ 1/2 & 1/2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -4 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1 \\ 5/2 \end{pmatrix}$$

1.5 *Lemma* A matrix changes bases if and only if it is nonsingular.

Proof For the 'only if' direction, if left-multiplication by a matrix changes bases then the matrix represents an invertible function, simply because we can invert the function by changing the bases back. Because it represents a function that is invertible, the matrix itself is invertible, and so is nonsingular.

1.5 *Lemma* A matrix changes bases if and only if it is nonsingular.

Proof For the 'only if' direction, if left-multiplication by a matrix changes bases then the matrix represents an invertible function, simply because we can invert the function by changing the bases back. Because it represents a function that is invertible, the matrix itself is invertible, and so is nonsingular.

For 'if' we will show that any nonsingular matrix M performs a change of basis operation from any given starting basis B (having n vectors, where the matrix is $n \times n$) to some ending basis.

If the matrix is the identity I then the statement is obvious. Otherwise because the matrix is nonsingular Corollary IV.**??** says there are elementary reduction matrices such that $R_r \cdots R_1 \cdot M = I$ with $r \geqslant 1$. Elementary matrices are invertible and their inverses are also elementary so multiplying both sides of that equation from the left by ${R_r}^{-1}$, then by ${R_{r-1}}^{-1}$, etc., gives M as a product of elementary matrices $M = {R_1}^{-1} \cdots {R_r}^{-1}$.

We will be done if we show that elementary matrices change a given basis to another basis, since then ${R_r}^{-1}$ changes B to some other basis B_r and ${R_{r-1}}^{-1}$ changes B_r to some B_{r-1}, etc. We will cover the three types of elementary matrices separately; recall the notation for the three.

$$M_i(k)\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ kc_i \\ \vdots \\ c_n \end{pmatrix} \quad P_{i,j}\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_j \\ \vdots \\ c_i \\ \vdots \\ c_n \end{pmatrix} \quad C_{i,j}(k)\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ kc_i + c_j \\ \vdots \\ c_n \end{pmatrix}$$

Applying a row-multiplication matrix $M_i(k)$ changes a representation with respect to $\langle\vec{\beta}_1,\ldots,\vec{\beta}_i,\ldots,\vec{\beta}_n\rangle$ to one with respect to $\langle\vec{\beta}_1,\ldots,(1/k)\vec{\beta}_i,\ldots,\vec{\beta}_n\rangle$.

$$\begin{aligned}\vec{v}=c_1\cdot\vec{\beta}_1+\cdots+c_i\cdot\vec{\beta}_i+\cdots+c_n\cdot\vec{\beta}_n&\\ \mapsto\ c_1\cdot\vec{\beta}_1+\cdots+kc_i\cdot(1/k)\vec{\beta}_i+\cdots+c_n\cdot\vec{\beta}_n=\vec{v}&\end{aligned}$$

The second one is a basis because the first is a basis and because of the $k\neq 0$ restriction in the definition of a row-multiplication matrix.

Applying a row-multiplication matrix $M_i(k)$ changes a representation with respect to $\langle\vec{\beta}_1,\ldots,\vec{\beta}_i,\ldots,\vec{\beta}_n\rangle$ to one with respect to $\langle\vec{\beta}_1,\ldots,(1/k)\vec{\beta}_i,\ldots,\vec{\beta}_n\rangle$.

$$\begin{aligned}\vec{v}=c_1\cdot\vec{\beta}_1+\cdots+c_i\cdot\vec{\beta}_i+\cdots+c_n\cdot\vec{\beta}_n&\\ \mapsto\; c_1\cdot\vec{\beta}_1+\cdots+kc_i\cdot(1/k)\vec{\beta}_i+\cdots+c_n\cdot\vec{\beta}_n=\vec{v}&\end{aligned}$$

The second one is a basis because the first is a basis and because of the $k\neq 0$ restriction in the definition of a row-multiplication matrix.

Similarly, left-multiplication by a row-swap matrix $P_{i,j}$ changes a representation with respect to the basis $\langle\vec{\beta}_1,\ldots,\vec{\beta}_i,\ldots,\vec{\beta}_j,\ldots,\vec{\beta}_n\rangle$ into one with respect to this basis $\langle\vec{\beta}_1,\ldots,\vec{\beta}_j,\ldots,\vec{\beta}_i,\ldots,\vec{\beta}_n\rangle$.

$$\begin{aligned}\vec{v}=c_1\cdot\vec{\beta}_1+\cdots+c_i\cdot\vec{\beta}_i+\cdots+c_j\vec{\beta}_j+\cdots+c_n\cdot\vec{\beta}_n&\\ \mapsto\; c_1\cdot\vec{\beta}_1+\cdots+c_j\cdot\vec{\beta}_j+\cdots+c_i\cdot\vec{\beta}_i+\cdots+c_n\cdot\vec{\beta}_n=\vec{v}&\end{aligned}$$

And, a representation with respect to $\langle\vec{\beta}_1,\ldots,\vec{\beta}_i,\ldots,\vec{\beta}_j,\ldots,\vec{\beta}_n\rangle$ changes via left-multiplication by a row-combination matrix $C_{i,j}(k)$ into a representation with respect to $\langle\vec{\beta}_1,\ldots,\vec{\beta}_i-k\vec{\beta}_j,\ldots,\vec{\beta}_j,\ldots,\vec{\beta}_n\rangle$

$$\begin{aligned}
\vec{v} &= c_1\cdot\vec{\beta}_1+\cdots+c_i\cdot\vec{\beta}_i+c_j\vec{\beta}_j+\cdots+c_n\cdot\vec{\beta}_n \\
\mapsto\ & c_1\cdot\vec{\beta}_1+\cdots+c_i\cdot(\vec{\beta}_i-k\vec{\beta}_j)+\cdots+(kc_i+c_j)\cdot\vec{\beta}_j+\cdots+c_n\cdot\vec{\beta}_n=\vec{v}
\end{aligned}$$

(the definition of $C_{i,j}(k)$ specifies that $i\neq j$ and $k\neq 0$). QED

1.6 *Corollary* A matrix is nonsingular if and only if it represents the identity map with respect to some pair of bases.

Changing map representations

The natural next step for us is to see how to convert $\mathrm{Rep}_{B,D}(h)$ to $\mathrm{Rep}_{\hat{B},\hat{D}}(h)$. Here is the arrow diagram.

$$\begin{array}{ccc} V_{\mathit{wrt}\ B} & \xrightarrow[H]{h} & W_{\mathit{wrt}\ D} \\ \mathrm{id}\big\downarrow & & \mathrm{id}\big\downarrow \\ V_{\mathit{wrt}\ \hat{B}} & \xrightarrow[\hat{H}]{h} & W_{\mathit{wrt}\ \hat{D}} \end{array}$$

To move from the lower-left to the lower-right we can either go straight over, or else up to V_B then over to W_D and then down. So we can calculate $\hat{H} = \mathrm{Rep}_{\hat{B},\hat{D}}(h)$ either by directly using $\hat{B}$ and $\hat{D}$, or else by first changing bases with $\mathrm{Rep}_{\hat{B},B}(\mathrm{id})$ then multiplying by $H = \mathrm{Rep}_{B,D}(h)$ and then changing bases with $\mathrm{Rep}_{D,\hat{D}}(\mathrm{id})$.

$$\hat{H} = \mathrm{Rep}_{D,\hat{D}}(\mathrm{id}) \cdot H \cdot \mathrm{Rep}_{\hat{B},B}(\mathrm{id}) \qquad (*)$$

Example Consider the derivative map $d/dx\colon \mathcal{P}_2 \to \mathcal{P}_2$, and consider also these two pairs of bases $B = \langle 1, 1+x, 1+x+x^2 \rangle, D = \langle 1+x^2, x, 1-x^2 \rangle$ and $B = \langle 1, x, x^2 \rangle, D = \langle 1+x, x+x^2, 1+x^2 \rangle$.

We can find H and $\hat{H}$ using the methods we have already seen.

$$\mathrm{Rep}_{B,D}(d/dx) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \mathrm{Rep}_{\hat{B},\hat{D}}(d/dx) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{pmatrix}$$

To do the conversion we find these.

$$\mathrm{Rep}_{\hat{B},B}(\mathrm{id}) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathrm{Rep}_{D,\hat{D}}(\mathrm{id}) = \begin{pmatrix} 0 & 1/2 & 1 \\ 0 & 1/2 & -1 \\ 1 & -1/2 & 0 \end{pmatrix}$$

Equation $(*)$ says that this equals $\mathrm{Rep}_{\hat{B},\hat{D}}(d/dx)$.

$$\begin{pmatrix} 0 & 1/2 & 1 \\ 0 & 1/2 & -1 \\ 1 & -1/2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

2.3 *Definition* Same-sized matrices H and $\hat{H}$ are *matrix equivalent* if there are nonsingular matrices P and Q such that $\hat{H} = PHQ$.

2.3 *Definition* Same-sized matrices H and $\hat{H}$ are *matrix equivalent* if there are nonsingular matrices P and Q such that $\hat{H} = PHQ$.

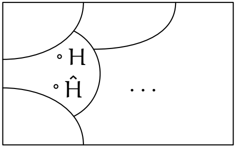
2.4 *Corollary* Matrix equivalent matrices represent the same map, with respect to appropriate pairs of bases.

2.3 *Definition* Same-sized matrices H and $\hat{H}$ are *matrix equivalent* if there are nonsingular matrices P and Q such that $\hat{H} = PHQ$.

2.4 *Corollary* Matrix equivalent matrices represent the same map, with respect to appropriate pairs of bases.

Exercise 19 checks that matrix equivalence is an equivalence relation. Thus it partitions the set of matrices into matrix equivalence classes.

All matrices:



H matrix equivalent to $\hat{H}$

Canonical form for matrix equivalence

2.6 *Theorem* Any $m \times n$ matrix of rank k is matrix equivalent to the $m \times n$ matrix that is all zeros except that the first k diagonal entries are ones.

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Canonical form for matrix equivalence

2.6 *Theorem* Any $m \times n$ matrix of rank k is matrix equivalent to the $m \times n$ matrix that is all zeros except that the first k diagonal entries are ones.

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

This is a *block partial-identity* form.

$$\left(\begin{array}{c|c} I & Z \\ \hline Z & Z \end{array}\right)$$

Proof Any $m \times n$ matrix of rank k is matrix equivalent to the $m \times n$ matrix that is all zeros except that the first k diagonal entries are ones.

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

QED

Matrix equivalence is characterized by rank

2.8 *Corollary* Matrix equivalence classes are characterized by rank: two same-sized matrices are matrix equivalent if and only if they have the same rank.

Matrix equivalence is characterized by rank

2.8 *Corollary* Matrix equivalence classes are characterized by rank: two same-sized matrices are matrix equivalent if and only if they have the same rank.

Proof Two same-sized matrices with the same rank are equivalent to the same block partial-identity matrix. QED

Matrix equivalence is characterized by rank

2.8 *Corollary* Matrix equivalence classes are characterized by rank: two same-sized matrices are matrix equivalent if and only if they have the same rank.

Proof Two same-sized matrices with the same rank are equivalent to the same block partial-identity matrix. QED

Example These two matrices are not matrix equivalent because Gauss's Method shows that the first has rank 3 while the second has rank 2.

$$\begin{pmatrix} 2 & 3 & 0 & -1 \\ 2 & 2 & 1 & 1 \\ 3 & 1 & 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & 5 & 1 & 4 \\ 2 & 0 & 5 & 1 \\ 3 & -5 & 9 & -2 \end{pmatrix}$$